

# Revisiting Optimal Eavesdropping in Quantum Cryptography: Choice of Interaction Is Unique up to Rotation of the Underlying Basis

Atanu Acharyya<sup>1</sup> and Goutam Paul<sup>2</sup>

<sup>1</sup>*Applied Statistics Unit,  
Indian Statistical Institute, Kolkata 700 108, India,  
Email: pub.academy.15@gmail.com*

<sup>2</sup>*Cryptology and Security Research Unit,  
R. C. Bose Centre for Cryptology and Security,  
Indian Statistical Institute, Kolkata 700 108, India,  
Email: goutam.paul@isical.ac.in*

A general framework of optimal eavesdropping on BB84 protocol was provided by Fuchs et al. [Phys. Rev. A, 1997]. An upper bound on mutual information was derived, which could be achieved by a specific type of interaction and the corresponding measurement. However, uniqueness of optimal interaction was posed as an unsolved problem there and it has remained open for almost two decades now. In this paper, we solve this open problem and establish the uniqueness of optimal interaction up to rotation. The specific choice of optimal interaction by Fuchs et al. is shown to be a special case of the form derived in our work.

## I. INTRODUCTION

Symmetric key cryptography requires a secret key to be shared or distributed between the sender (say, Alice) and the receiver (say, Bob). The security of classical key distribution is based on hardness assumptions for solving certain computational problems. This gives security for computationally bounded adversary in the classical domain, but fails to guarantee security against quantum attacks. Quantum key distribution (QKD) is based on the principles of quantum mechanics. To encode classical bits, QKD uses quantum states which the attacker (say, Eve) cannot measure without creating disturbance detectable by Bob. QKD protocol does not require any computation complexity assumption and is provably secure against both classical as well as quantum adversaries.

The first and possibly the most celebrated QKD protocol is BB84 [1]. The protocol relies on the use of orthogonal states from one of the two conjugate bases, say,  $x$ - $y$  and  $u$ - $v$ , to encode a bit-string in qubits (e.g., polarized photons). Alice randomly selects one of the two bases and encodes 0 and 1 respectively by a qubit prepared in one of the two states in each base. Say, Alice encodes 0 to  $|x\rangle$  or  $|u\rangle$ , and 1 to  $|y\rangle$  or  $|v\rangle$ , depending on the chosen basis. When Bob receives a state from Alice, he randomly selects a basis  $x$ - $y$  or  $u$ - $v$  and makes a measurement. Once the measurement is done for all the received qubits, Alice and Bob publicly announce the sequence of bases used by them and discard the bits where the bases do not match. The resulting bit string, followed by error correction and privacy amplification, becomes the common secret key. However, presence of an eavesdropper may disturb the state of a qubit sent by Alice for which Bob may get a wrong result even if the corresponding bases of measurement between Alice and Bob match. To overcome this problem, Alice and Bob sacrifice some of the bits by comparing their values publicly.

Fuchs et al. [2] provided a general framework of optimal eavesdropping on BB84 protocol. The authors derived an upper bound on mutual information, described a specific type of interaction and the corresponding measurement that achieves the bound. They finally explained an optimal strategy for Eve in interpreting her measurement. However, the optimal interaction described there was a specific choice and the uniqueness of the optimal interaction was left as an open problem. They commented:

*“It is easy to check that the solution here is correct, but the extent to which it is unique aside from trivial changes of basis and of phase remains unknown.”*

Interestingly, this problem has been open for last two decades. In this paper, we solve this open problem and establish the uniqueness. Here we do not claim any improvement of optimal information gain over [2]. We characterize the classes of interaction that can achieve the already-existing optimal bound given by [2]. We have shown that the choice of optimal interaction in [2] is a special case of the generalized form provided by us. We also explicitly show the corresponding optimal measurement by Eve.

The content of this paper is organized as follows. Section II explains basic terminologies used for optimal eavesdropping introduced in [2]. Section III contains summary of certain results from [2] which are relevant to our work. Our results are explained in Section IV. The remaining portion discusses the connection of our results with [2] followed by a conclusion.

## II. PRELIMINARIES

Alice and Bob want to share a secret key using BB84 protocol. Alice randomly chooses a basis from  $\mathfrak{B}_{xy} =$

$\{|x\rangle, |y\rangle\}$  and  $\mathfrak{B}_{uv} = \{|u\rangle, |v\rangle\}$ , where

$$|x\rangle = \frac{1}{\sqrt{2}}(|u\rangle + |v\rangle), \quad |y\rangle = \frac{1}{\sqrt{2}}(|u\rangle - |v\rangle), \quad (1)$$

i.e., the bases are conjugate to each other. Note that a more common notation uses  $|0\rangle, |1\rangle, |+\rangle$  and  $|-\rangle$  instead of  $|x\rangle, |y\rangle, |u\rangle$  and  $|v\rangle$  respectively. However, we follow the same notations as in Fuchs et al. [2] so that the connection to their work is easily visible.

Alice encodes her key-bits, each as a polarized photon, and sends it to Bob. Suppose, an eavesdropper Eve interferes the communication while she lets a probe interact unitarily with the qubit sent by Alice.

Suppose Alice has chosen a signal, say,  $|x\rangle$  (corresponding density operator being  $\rho_x^A = |x\rangle\langle x|$ ), in the basis  $\mathfrak{B}_{xy}$ . Eve lets a probe, initially in state  $|\psi_0\rangle$  (corresponding density operator  $\rho_0^E = |\psi_0\rangle\langle\psi_0|$ ), interact unitarily (realized by a unitary operator  $\mathcal{U}$ ) with the qubit sent by Alice. The post-interaction joint state  $|X\rangle$  between Alice and Eve, which is an entangled state of the probe of Eve and the photon sent by Alice, is realized by

$$|x\rangle \otimes |\psi_0\rangle \xrightarrow{\mathcal{U}} |X\rangle.$$

Bob receives a simple mixture of the two basis vectors (here  $\mathfrak{B}_{xy}$ ) chosen by Alice, i.e., Bob's density matrix is always diagonal in the basis chosen by Alice. Thus, Schmidt decomposition of the post-interaction joint state  $|X\rangle$  must be of the form

$$|X\rangle = \sqrt{\alpha} |x\rangle |\xi_x\rangle + \sqrt{1-\alpha} |y\rangle |\zeta_x\rangle,$$

such that

$$|\xi_x\rangle \perp |\zeta_x\rangle, \quad (2)$$

where  $|\xi_x\rangle, |\zeta_x\rangle$  are component of Eve's part of the joint state after the interaction.

Similarly, when Alice sends  $|y\rangle$ , the post-interaction state  $|Y\rangle$  must be of the form

$$|Y\rangle = \sqrt{\beta} |y\rangle |\xi_y\rangle + \sqrt{1-\beta} |x\rangle |\zeta_y\rangle,$$

such that

$$|\xi_y\rangle \perp |\zeta_y\rangle. \quad (3)$$

The density operator for the post-interaction state  $|X\rangle$  is given by

$$\rho_x^{AE} = |X\rangle\langle X| = \mathcal{U} (\rho_x^A \otimes \rho_0^E) \mathcal{U}^\dagger. \quad (4)$$

Eve's description of the system will be

$$\rho_x := \rho_x^E = \text{tr}_A (\rho_x^{AE}) = \text{tr}_A (|X\rangle\langle X|), \quad (5)$$

where  $\text{tr}_A$  represents partial trace over Alice's qubit.

Since the interaction is unitary, it follows from Equations (1) and (4) that

$$|X\rangle = \frac{1}{\sqrt{2}}(|U\rangle + |V\rangle), \quad |Y\rangle = \frac{1}{\sqrt{2}}(|U\rangle - |V\rangle). \quad (6)$$

Before performing any measurement, Eve waits until Alice declares her choice of basis publicly. Eve's measurement is considered to be a Positive Operator-Valued Measure (POVM)  $\{E_\lambda\}$  or  $\{F_\lambda\}$  depending on whether Alice's choice is  $x$ - $y$  or  $u$ - $v$  basis. Note that the operators  $\{E_\lambda\}$  satisfy two properties [3, 4]: they are all non-negative definite, i.e.,

$$\langle \gamma | E_\lambda | \gamma \rangle \geq 0, \quad \forall |\gamma\rangle,$$

and satisfy the completeness relation

$$\sum_\lambda E_\lambda = \mathbb{1}.$$

Suppose, Alice sends a signal in  $x$ - $y$  (or,  $u$ - $v$ ) basis with the prior probabilities  $p_x, p_y$  (or,  $p_u, p_v$ ) respectively. Once Alice reveals her basis to be  $x$ - $y$ , Eve uses a POVM  $\{E_\lambda\}$  to perform a measurement on her probe. Considering  $\mathcal{A}, \mathcal{B}, \mathcal{E}$  as random variables corresponding to the signal sent by Alice, signal received by Bob, and, measurement outcome of Eve, the conditional probability of the various outcomes  $\lambda$  of that measurement is given by

$$P_{\lambda x} := \Pr[\mathcal{E} = \lambda | \mathcal{A} = x] = \text{tr}(\rho_x E_\lambda), \quad (7)$$

$$P_{\lambda y} := \Pr[\mathcal{E} = \lambda | \mathcal{A} = y] = \text{tr}(\rho_y E_\lambda). \quad (8)$$

Henceforth, we use the notation  $:=$  to denote "defined as". The probability that Eve gets outcome  $\lambda$ , when Alice uses  $x$ - $y$  basis is thus

$$q_{xy}(\lambda) := \Pr[\mathcal{E} = \lambda] = P_{\lambda x} p_x + P_{\lambda y} p_y.$$

Looking at outcome  $\lambda$ , Eve assigns a guess for the signal sent by Alice following some strategy. The posterior probability  $Q_{x\lambda}$  (or  $Q_{y\lambda}$ ) of the event that Alice had sent  $x$  (or  $y$ ) given that Eve has observed  $\lambda$  is given by Bayes' theorem.

$$Q_{x\lambda} := \Pr[\mathcal{A} = x | \mathcal{E} = \lambda] = \frac{P_{\lambda x} p_x}{q_{xy}(\lambda)},$$

$$Q_{y\lambda} := \Pr[\mathcal{A} = y | \mathcal{E} = \lambda] = \frac{P_{\lambda y} p_y}{q_{xy}(\lambda)}.$$

A simple way that Eve can utilize these likelihoods is to perform a guess given by the following function.

$$\text{argmax} \{Q_{x\lambda}, Q_{y\lambda}\} = \begin{cases} x, & \text{if } Q_{x\lambda} > Q_{y\lambda}, \\ y, & \text{if } Q_{y\lambda} > Q_{x\lambda}. \end{cases}$$

A convenient measure of Eve's **information gain** for an outcome  $\lambda$ , as proposed in [2], is

$$G_{xy}(\lambda) := |Q_{x\lambda} - Q_{y\lambda}|.$$

On average, Eve's information gain over all outcomes is

$$G_{xy} := \sum_\lambda q_{xy}(\lambda) G_{xy}(\lambda) = \sum_\lambda |P_{\lambda x} p_x - P_{\lambda y} p_y|.$$

In particular, for equiprobable signals,

$$G_{xy} = \frac{1}{2} \sum_{\lambda} |P_{\lambda x} - P_{\lambda y}|.$$

A more sophisticated data processing by Eve is **mutual information** [5]. For equal prior, this is given by

$$I_{xy} := \ln 2 + \sum_{\lambda} q_{xy}(\lambda) (Q_{x\lambda} \ln Q_{x\lambda} + Q_{y\lambda} \ln Q_{y\lambda}).$$

Eve's attempt to measure the probe creates **disturbance** to the signal sent by Alice which is detectable by Bob. For signal sent in  $x$ - $y$  basis, the disturbance incorporated by Eve could be described by

$$D_{xy} := \sum_{\lambda} q_{xy}(\lambda) d_{xy}(\lambda),$$

where,  $d_{xy}(\lambda)$  is the avg error for Bob to read the signal sent by Alice while Eve detects  $\lambda$ . For equal prior,

$$d_{xy}(\lambda) := \frac{1}{2} (d_{\lambda x} + d_{\lambda y}),$$

where,  $d_{\lambda x}$  is the error for Bob when Alice sends  $x$  while Eve detects  $\lambda$  (i.e., Bob reads  $y$ ), i.e.,

$$d_{\lambda x} := \Pr[\mathcal{B} = y | (\mathcal{A} = x, \mathcal{E} = \lambda)],$$

and  $d_{\lambda y}$  is the error for Bob when Alice sends  $y$  while Eve detects  $\lambda$  (i.e., Bob reads  $x$ ), i.e.,

$$d_{\lambda y} := \Pr[\mathcal{B} = x | (\mathcal{A} = y, \mathcal{E} = \lambda)].$$

Clearly,  $D_{xy}$  is the observable error rate of Bob to read the signal sent by Alice prepared in  $x$ - $y$  basis.

Similarly, one can define  $G_{uv}$ ,  $I_{uv}$ ,  $D_{uv}$  while considering Alice's signal was prepared in  $u$ - $v$  basis.

### III. SUMMARY OF OPTIMAL EAVESDROPPING BY FUCHS ET AL. [2]

In this section, we recollect the results given by [2]: an upper bound on information gain ( $G$ ) and mutual information ( $I$ ), followed by a necessary and sufficient condition to achieve the bounds and finally an optimal interaction (and the corresponding POVM) for unequal and equal error rates. For any quantity  $q$ , we denote its optimal (maximum) value by  $q^*$ .

#### A. Upper Bounds on Information Gain ( $G$ ) and Mutual Information ( $I$ )

Fuchs et al. [2] provided an upper bound on the information gain ( $G$ ). This bound was used to provide with an upper bound on the mutual information ( $I$ ). A necessary and sufficient condition to achieve the bounds was given there. We recollect these results here.

**Proposition 1. (An upper bound on information gain ( $G$ ))** [2, Equation (23,24)]  
For a given POVM  $\{E_{\lambda}\}$ ,

$$G_{xy} \leq 2\sqrt{D_{uv}(1-D_{uv})}, \quad (9)$$

$$G_{uv} \leq 2\sqrt{D_{xy}(1-D_{xy})}. \quad (10)$$

Moreover, for measurement outcome  $\lambda$  of Eve, the bound on information gain [2, Equation (20)] can be expressed by the following inequality

$$G_{xy}(\lambda) \leq 2\sqrt{d_{uv}(\lambda)(1-d_{uv}(\lambda))}. \quad (11)$$

It is interesting to note that while Eve's information gain refers to signals sent in the  $x$ - $y$  basis, Bob's error rate refers to signals sent in the  $u$ - $v$  basis and vice versa.

**Proposition 2. (An Upper Bound on Mutual Information ( $I$ ))** [2, Equation (31,32)]  
For a given POVM  $\{E_{\lambda}\}$ ,

$$I_{xy} \leq \frac{1}{2} \phi\left(2\sqrt{D_{uv}(1-D_{uv})}\right), \quad (12)$$

$$I_{uv} \leq \frac{1}{2} \phi\left(2\sqrt{D_{xy}(1-D_{xy})}\right), \quad (13)$$

where  $\phi(z) = (1+z)\ln(1+z) + (1-z)\ln(1-z)$ .

Subscripts emphasize that the mutual information and error rate refer to signals sent in two different bases.

**Proposition 3. (Necessary and Sufficient Conditions to Achieve  $G^*$ )** [2, Equation (38,39)]  
The necessary and sufficient conditions for equality in Equation (9) are

$$|V_{\lambda u}\rangle = \varepsilon_{\lambda} \sqrt{\frac{D_{uv}}{1-D_{uv}}} |U_{\lambda u}\rangle \quad (14)$$

and

$$|U_{\lambda v}\rangle = \varepsilon_{\lambda} \sqrt{\frac{D_{uv}}{1-D_{uv}}} |V_{\lambda v}\rangle, \quad (15)$$

where

$$\varepsilon_{\lambda} = \pm 1 = \text{sgn}(Q_{x\lambda} - Q_{y\lambda}) \quad (16)$$

and

$$\begin{aligned} |U_{\lambda u}\rangle &= B_u \otimes \sqrt{E_{\lambda}} |U\rangle, & |V_{\lambda u}\rangle &= B_u \otimes \sqrt{E_{\lambda}} |V\rangle, \\ |U_{\lambda v}\rangle &= B_v \otimes \sqrt{E_{\lambda}} |U\rangle, & |V_{\lambda v}\rangle &= B_v \otimes \sqrt{E_{\lambda}} |V\rangle, \\ B_u &= |u\rangle\langle u|, & B_v &= |v\rangle\langle v|. \end{aligned} \quad (17)$$

It is intriguing to note that the set of conditions that optimize  $G$  also optimize  $I$ . Therefore, the necessary and sufficient conditions for equality in Equation (12) is also the same as those in Proposition 3. Hence, any measurement or strategy that maximizes  $G$  also maximizes  $I$ .

### B. Description of the Post-interaction States

$|X\rangle, |Y\rangle$

For optimal  $G_{xy}$ , the post-interaction states are

$$\begin{aligned} |X\rangle &= \sqrt{1-D_{xy}} |x\rangle|\xi_x\rangle + \sqrt{D_{xy}} |y\rangle|\zeta_x\rangle, \\ |Y\rangle &= \sqrt{1-D_{xy}} |y\rangle|\xi_y\rangle + \sqrt{D_{xy}} |x\rangle|\zeta_y\rangle. \end{aligned} \quad (18)$$

Assuming that all inner products  $\langle\xi_i|\zeta_j\rangle$  are real, the restriction on  $|\xi_i\rangle, |\zeta_j\rangle$  in Equations (2) and (3) becomes more restricted as

$$\{|\xi_x\rangle, |\xi_y\rangle\} \perp \{|\zeta_x\rangle, |\zeta_y\rangle\}. \quad (19)$$

Similarly, for optimal  $G_{uv}$ , the post-interaction states are

$$\begin{aligned} |U\rangle &= \sqrt{1-D_{uv}} |u\rangle|\xi_u\rangle + \sqrt{D_{uv}} |v\rangle|\zeta_u\rangle, \\ |V\rangle &= \sqrt{1-D_{uv}} |v\rangle|\xi_v\rangle + \sqrt{D_{uv}} |u\rangle|\zeta_v\rangle. \end{aligned} \quad (20)$$

Since the bases  $\mathfrak{B}_{xy}$  and  $\mathfrak{B}_{uv}$  are conjugate to each other, we expect to get a relationship between  $|\xi_i\rangle, |\zeta_j\rangle$  in  $u-v$  basis and those in  $x-y$  basis which is described below.

$$\begin{aligned} &2\sqrt{1-D_{uv}} |\xi_u\rangle \\ &= \sqrt{1-D_{xy}} (|\xi_x\rangle + |\xi_y\rangle) + \sqrt{D_{xy}} (|\zeta_x\rangle + |\zeta_y\rangle), \\ &2\sqrt{D_{uv}} |\zeta_u\rangle \\ &= \sqrt{1-D_{xy}} (|\xi_x\rangle - |\xi_y\rangle) + \sqrt{D_{xy}} (|\zeta_y\rangle - |\zeta_x\rangle). \end{aligned} \quad (21)$$

Similarly,

$$\begin{aligned} &2\sqrt{1-D_{uv}} |\xi_v\rangle \\ &= \sqrt{1-D_{xy}} (|\xi_x\rangle + |\xi_y\rangle) - \sqrt{D_{xy}} (|\zeta_x\rangle + |\zeta_y\rangle), \\ &2\sqrt{D_{uv}} |\zeta_v\rangle \\ &= \sqrt{1-D_{xy}} (|\xi_x\rangle - |\xi_y\rangle) - \sqrt{D_{xy}} (|\zeta_y\rangle - |\zeta_x\rangle). \end{aligned} \quad (22)$$

From the orthogonality relation (19), one can say that Eve's probe lives in a Hilbert space of at most 4-dimensions, and thus is taken to be made of 2 qubits (4 states). It is therefore convenient to introduce same bases ( $x-y$  and  $u-v$ , used by Alice) for each of Eve's qubits.

### C. Optimal Interaction to maximize $G, I$ : a Specific Choice

Any interaction, as described above, that leads to optimality could be chosen. In [2, Section III: Equations (50,51)], one such specific choice is made for unequal error rates, which is shown to be a correct choice (correct in the sense that the choice leads to optimality). Similarly, for equal error rates, another specific choice is made in [2, Section IV, Equation (69)]. However, uniqueness of the choice was left as an open problem in [2, Section III, first paragraph].

#### 1. For Unequal Error Rates, i.e., $D_{xy} \neq D_{uv}$

Equations (50), (51) of [2, Section III] are restated here. Consider a canonical basis for Eve's probe as  $\{|\mathcal{E}_0\rangle, |\mathcal{E}_1\rangle, |\mathcal{E}_2\rangle, |\mathcal{E}_3\rangle\}$ . Without loss of generality (w.l.o.g.),

$$|\mathcal{E}_0\rangle = |x\rangle|x\rangle, |\mathcal{E}_1\rangle = |y\rangle|x\rangle, |\mathcal{E}_2\rangle = |x\rangle|y\rangle, |\mathcal{E}_3\rangle = |y\rangle|y\rangle. \quad (23)$$

To describe  $|\xi_i\rangle, |\zeta_j\rangle$ , the work [2] considered an orthonormal set, namely, the Bell Basis with respect to (w.r.t.)  $x-y$ , as follows.

$$\begin{aligned} |\Phi_{xy}^\pm\rangle &:= \frac{1}{\sqrt{2}} (|x\rangle|x\rangle \pm |y\rangle|y\rangle) = \frac{1}{\sqrt{2}} (|\mathcal{E}_0\rangle \pm |\mathcal{E}_3\rangle), \\ |\Psi_{xy}^\pm\rangle &:= \frac{1}{\sqrt{2}} (|x\rangle|y\rangle \pm |y\rangle|x\rangle) = \frac{1}{\sqrt{2}} (|\mathcal{E}_2\rangle \pm |\mathcal{E}_1\rangle). \end{aligned} \quad (24)$$

In terms of the Bell basis vectors for Eve's probe, the interaction was chosen such that

$$\begin{aligned} |\xi_x\rangle &= \sqrt{1-D_{uv}} |\Phi_{xy}^+\rangle + \sqrt{D_{uv}} |\Phi_{xy}^-\rangle, \\ |\xi_y\rangle &= \sqrt{1-D_{uv}} |\Phi_{xy}^+\rangle - \sqrt{D_{uv}} |\Phi_{xy}^-\rangle, \\ |\zeta_x\rangle &= \sqrt{1-D_{uv}} |\Psi_{xy}^+\rangle - \sqrt{D_{uv}} |\Psi_{xy}^-\rangle, \\ |\zeta_y\rangle &= \sqrt{1-D_{uv}} |\Psi_{xy}^+\rangle + \sqrt{D_{uv}} |\Psi_{xy}^-\rangle. \end{aligned} \quad (25)$$

The corresponding optimal POVM, as shown in [2, Equations (55,56)], is described below.

$$E_\lambda = |E_\lambda\rangle\langle E_\lambda|, \quad (26)$$

where

$$|E_0\rangle = |\mathcal{E}_0\rangle, |E_1\rangle = |\mathcal{E}_1\rangle, |E_2\rangle = |\mathcal{E}_2\rangle, |E_3\rangle = |\mathcal{E}_3\rangle.$$

Introducing new notations  $\mathcal{D}_{uv}, \bar{\mathcal{D}}_{uv}$ , we can write a closed form of  $|\xi_i\rangle, |\zeta_j\rangle$  as below.

$$\begin{aligned} |\xi_x\rangle &= \mathcal{D}_{uv} |\mathcal{E}_0\rangle + \bar{\mathcal{D}}_{uv} |\mathcal{E}_3\rangle, \\ |\xi_y\rangle &= \bar{\mathcal{D}}_{uv} |\mathcal{E}_0\rangle + \mathcal{D}_{uv} |\mathcal{E}_3\rangle, \\ |\zeta_x\rangle &= \bar{\mathcal{D}}_{uv} |\mathcal{E}_2\rangle + \mathcal{D}_{uv} |\mathcal{E}_1\rangle, \\ |\zeta_y\rangle &= \mathcal{D}_{uv} |\mathcal{E}_2\rangle + \bar{\mathcal{D}}_{uv} |\mathcal{E}_1\rangle, \end{aligned} \quad (27)$$

where

$$\begin{aligned} \mathcal{D}_{uv} &:= \frac{\sqrt{1-D_{uv}} + \sqrt{D_{uv}}}{\sqrt{2}}, \\ \bar{\mathcal{D}}_{uv} &:= \frac{\sqrt{1-D_{uv}} - \sqrt{D_{uv}}}{\sqrt{2}}. \end{aligned} \quad (28)$$

The following relations appear to be useful.

$$\begin{aligned} \mathcal{D}_{uv} \cdot \bar{\mathcal{D}}_{uv} &= \frac{1}{2} (1 - 2D_{uv}), \\ \mathcal{D}_{uv}^2 + \bar{\mathcal{D}}_{uv}^2 &= 1, \\ \mathcal{D}_{uv}^2 - \bar{\mathcal{D}}_{uv}^2 &= 2\sqrt{D_{uv}(1-D_{uv})}. \end{aligned} \quad (29)$$

2. For Equal Error Rates, i.e.,  $D_{xy} = D_{uv} = D$

For equal error rates, [2, Section IV, Equation (69)] comes up with another choice of  $|\xi_i\rangle, |\zeta_j\rangle$ . We describe it as below.

$$\begin{aligned} |\xi_x\rangle &= |x\rangle|x\rangle, \\ |\xi_y\rangle &= (\cos\alpha|x\rangle + \sin\alpha|y\rangle)|x\rangle \\ |\zeta_x\rangle &= |x\rangle|y\rangle, \\ |\zeta_y\rangle &= (\cos\beta|x\rangle + \sin\beta|y\rangle)|y\rangle. \end{aligned} \quad (30)$$

Optimality of  $G$  (or  $I$ ) is reached when

$$\alpha = \beta \quad \text{and} \quad \sin\alpha = 2\sqrt{D(1-D)} = \mathcal{D}^2 - \overline{\mathcal{D}}^2.$$

The notations  $\mathcal{D}, \overline{\mathcal{D}}$  have already been defined in Equations (28).

Thus, the optimal interaction can be written as

$$\begin{aligned} |\xi_x\rangle &= |\mathcal{E}_0\rangle, \\ |\xi_y\rangle &= 2\mathcal{D} \cdot \overline{\mathcal{D}} |\mathcal{E}_0\rangle + (\mathcal{D}^2 - \overline{\mathcal{D}}^2) |\mathcal{E}_1\rangle, \\ |\zeta_x\rangle &= |\mathcal{E}_2\rangle, \\ |\zeta_y\rangle &= 2\mathcal{D} \cdot \overline{\mathcal{D}} |\mathcal{E}_2\rangle + (\mathcal{D}^2 - \overline{\mathcal{D}}^2) |\mathcal{E}_3\rangle. \end{aligned} \quad (31)$$

However, the corresponding optimal POVM was not shown explicitly in [2], which we establish in Section IV C.

#### IV. OUR RESULTS

In this section, we show that the interaction by Eve (and the corresponding POVM) that leads to optimal information gain has a unique form in a fixed basis.

##### A. Optimal Measurement (POVM) to Maximize Information Gain ( $G$ ): a Generic Condition

Let's consider the problem below:

$$\text{maximize } G_{xy} := \sum_{\lambda} |P_{\lambda x} p_x - P_{\lambda y} p_y|$$

over all POVM  $\{E_{\lambda}\}$ .

In [3], an optimal measurement for this maximization was derived. There, the maximization was done on *Kolmogorov Variational Distance* [3, Equation (130)]. The calculation was performed in [3, Appendix (Section 7)], which shows that the optimal measurement corresponds to a Hermitian operator given by [3, Equation (21)] and the optimal POVM is an orthonormal eigenbasis of that operator. We describe the result here with a proof in terms of maximizing  $G$ . Note that this result is presented here for the sake of completeness and easy reference and we do not claim any contribution for this result.

**Lemma 1.** An optimal POVM to attain  $G_{xy}^*$  is an orthonormal eigenprojector  $\{E_{\lambda}\}$  that diagonalize the Hermitian operator

$$\tilde{\Gamma}_{xy} := p_x \rho_x - p_y \rho_y, \quad (32)$$

where  $\rho_x$ , as defined in Equation (5), is the partial trace (over Alice's qubit) of the post-interaction state  $|X\rangle$ .

*Proof.*

$$\begin{aligned} G_{xy} &:= \sum_{\lambda} |P_{\lambda x} p_x - P_{\lambda y} p_y| \\ &= \sum_{\lambda} |p_x \text{tr}(\rho_x E_{\lambda}) - p_y \text{tr}(\rho_y E_{\lambda})|, \text{ using Eq. (7)} \\ &= \sum_{\lambda} \left| \text{tr}(\tilde{\Gamma} E_{\lambda}) \right|, \text{ using Eq. (32)} \\ &= \sum_{\lambda} \left| \sum_i \gamma_i \langle \gamma_i | E_{\lambda} | \gamma_i \rangle \right| \\ &\quad [\text{where } \{|\gamma_i\rangle\}_i \text{ is a normalized eigenbasis for } \tilde{\Gamma} \\ &\quad \text{and } \{\gamma_i\}_i \text{ are the associated eigenvalues}] \\ &\leq \sum_{\lambda} \sum_i |\gamma_i| \langle \gamma_i | E_{\lambda} | \gamma_i \rangle \\ &\quad [\text{equality occurs for orthogonal } \{|\gamma_i\rangle\}_i] \\ &= \sum_i |\gamma_i| \langle \gamma_i | \sum_{\lambda} E_{\lambda} | \gamma_i \rangle \\ &= \sum_i |\gamma_i| = \text{tr} |\tilde{\Gamma}_{xy}|. \end{aligned}$$

Thus,  $G_{xy}^*$  is achieved by some POVM  $\{E_{\lambda}\}$  that are projectors onto an orthonormal eigenbasis of  $\tilde{\Gamma}_{xy}$ .  $\square$

**Remark 1.** Since we consider equal prior probabilities here, analogous to Equation (32), we define

$$\Gamma_{xy} := \rho_x - \rho_y \quad (33)$$

and use it throughout the rest of the paper.

We draw a basic flowchart in Figure 1 that inter-relates optimal interaction and corresponding POVM via the matrix  $\Gamma_{xy}$ .

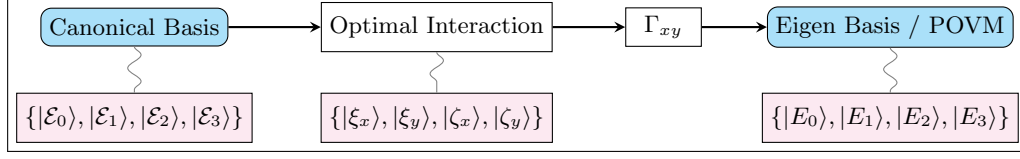
##### B. Optimal Interaction to Maximize Information Gain ( $G$ ): a Generic Form of Optimal $|\xi_i\rangle, |\zeta_j\rangle$

We use the following result for equal priors to find an expression of  $|\xi_i\rangle, |\zeta_j\rangle$  for optimal interaction.

**Lemma 2.** Optimality conditions for  $G_{xy}$  ensures that each  $G_{xy}^*(\lambda)$  is equal to  $G_{xy}^*$  and the corresponding optimal value is given by

$$G_{xy}^* = 2\sqrt{D_{uv}(1-D_{uv})} = G_{xy}^*(\lambda), \quad \forall \lambda. \quad (34)$$

FIG. 1: Optimal interaction to POVM



*Proof.* For signal sent in  $x$ - $y$  basis, the optimal information gain, by Equation (9), is

$$G_{xy}^* = 2\sqrt{D_{uv}(1 - D_{uv})}. \quad (35)$$

By Equation (11), for measurement outcome  $\lambda$  of Eve,

$$G_{xy}^*(\lambda) = 2\sqrt{d_{uv}(\lambda)(1 - d_{uv}(\lambda))} \quad (36)$$

For optimality, the necessary and sufficient conditions in Proposition 3 must be satisfied. By [2, Equation (20)], this requires

$$d_{uv}(\lambda) = D_{uv}, \quad \forall \lambda \quad (37)$$

which ensures that the lemma is proved.  $\square$

**Note 1.** Since we consider equal prior probabilities, we use the following working formula of  $G_{xy}(\lambda)$  in the process of the derivation,

$$G_{xy}(\lambda) := |Q_{x\lambda} - Q_{y\lambda}| = \frac{|P_{\lambda x} - P_{\lambda y}|}{P_{\lambda x} + P_{\lambda y}}. \quad (38)$$

Here we describe an expression of  $P_{\lambda x}, P_{\lambda y}$  in terms of  $|\xi_i\rangle, |\zeta_j\rangle$  and a POVM  $\{E_\lambda\}$ .

**Theorem 1.** For a POVM  $\{E_\lambda\}_{\lambda \in \{0,1,2,3\}}$

$$P_{\lambda x} = (1 - D_{xy})\langle \xi_x | E_\lambda \rangle^2 + D_{xy}\langle \zeta_x | E_\lambda \rangle^2, \quad (39)$$

$$P_{\lambda y} = (1 - D_{xy})\langle \xi_y | E_\lambda \rangle^2 + D_{xy}\langle \zeta_y | E_\lambda \rangle^2. \quad (40)$$

*Proof.* Using Equation (18) in Equation (5), we get,

$$\rho_x := \text{Tr}_A(|X\rangle\langle X|) = (1 - D_{xy})\hat{\xi}_x + D_{xy}\hat{\zeta}_x, \quad (41)$$

where

$$\hat{\xi}_x := |\xi_x\rangle\langle \xi_x|, \quad \hat{\zeta}_x := |\zeta_x\rangle\langle \zeta_x|.$$

By Equation (7),

$$\begin{aligned} P_{\lambda x} &= \text{Tr}(\rho_x E_\lambda) \\ &= (1 - D_{xy})\text{Tr}(\hat{\xi}_x E_\lambda) + D_{xy}\text{Tr}(\hat{\zeta}_x E_\lambda) \\ &= (1 - D_{xy})\langle \xi_x | E_\lambda \rangle^2 + D_{xy}\langle \zeta_x | E_\lambda \rangle^2. \end{aligned}$$

Similarly, we can derive an expression for  $P_{\lambda y}$ .  $\square$

Now, consider an arbitrary orthonormal eigenbasis of  $\Gamma_{xy}$  (which corresponds to an optimal POVM that leads to optimal  $G_{xy}$ ). We derive the general form of  $|\xi_i\rangle, |\zeta_j\rangle$ , described in that eigenbasis, for optimal interaction.

**Theorem 2.** Let  $\{|E_\lambda\rangle\}$  be an orthonormal eigenbasis of  $\Gamma_{xy}$ , the corresponding eigenprojectors being  $\{E_\lambda\}$ . Then for optimal interaction, the general form of  $|\xi_i\rangle, |\zeta_j\rangle$ , described in that eigenbasis becomes

$$\begin{aligned} |\xi_x\rangle &= \mathcal{D}_{uv} |E_0\rangle + \overline{\mathcal{D}}_{uv} |E_1\rangle, \\ |\xi_y\rangle &= \overline{\mathcal{D}}_{uv} |E_0\rangle + \mathcal{D}_{uv} |E_1\rangle, \\ |\zeta_x\rangle &= \mathcal{D}_{uv} |E_2\rangle + \overline{\mathcal{D}}_{uv} |E_3\rangle, \\ |\zeta_y\rangle &= \overline{\mathcal{D}}_{uv} |E_2\rangle + \mathcal{D}_{uv} |E_3\rangle, \end{aligned} \quad (42)$$

where  $\mathcal{D}_{uv}, \overline{\mathcal{D}}_{uv}$  are as defined in Equation (28).

*Proof.* First we need to fix an orthonormal basis to describe  $|\xi_i\rangle, |\zeta_j\rangle$  following restriction (19). For that purpose, there is no harm to choose the above eigenbasis to describe  $|\xi_i\rangle, |\zeta_j\rangle$ . Orthogonality restriction (19) is automatically satisfied if we choose  $|\xi_i\rangle \in \text{span}\{|E_0\rangle, |E_1\rangle\}$  and  $|\zeta_j\rangle \in \text{span}\{|E_2\rangle, |E_3\rangle\}$ . So the general form of  $|\xi_i\rangle, |\zeta_j\rangle$  becomes

$$\begin{aligned} |\xi_x\rangle &= \sqrt{\alpha} |E_0\rangle + \sqrt{1 - \alpha} |E_1\rangle, \\ |\xi_y\rangle &= \sqrt{\beta} |E_0\rangle + \sqrt{1 - \beta} |E_1\rangle, \\ |\zeta_x\rangle &= \sqrt{\mu} |E_2\rangle + \sqrt{1 - \mu} |E_3\rangle, \\ |\zeta_y\rangle &= \sqrt{\nu} |E_2\rangle + \sqrt{1 - \nu} |E_3\rangle. \end{aligned} \quad (43)$$

Using this form of  $|\xi_i\rangle, |\zeta_j\rangle$  in Equation (39), we find values of  $G_{xy}(\lambda)$  as shown in Table I.

By Lemma 2, for optimal  $G_{xy}$ ,  $G_{xy}(\lambda)$ 's are equal. Equating  $G_{xy}(0), G_{xy}(1)$  in Table I, we get,

$$\alpha + \beta = 1, \quad G_{xy}(0) = G_{xy}(1) = |2\alpha - 1|.$$

Similarly, equating  $G_{xy}(2), G_{xy}(3)$  in Table I, we get,

$$\mu + \nu = 1, \quad G_{xy}(2) = G_{xy}(3) = |2\mu - 1|.$$

Together, equating  $G_{xy}(0), G_{xy}(2)$ , we get,

$$\mu = \alpha, \quad \nu = \beta = 1 - \alpha. \quad (44)$$

Thus,

$$G_{xy}^*(0) = \mathcal{D}_{uv}^2 - \overline{\mathcal{D}}_{uv}^2 = 2\mathcal{D}_{uv}^2 - 1 = |2\alpha - 1|$$

gives rise to

$$\sqrt{\alpha} = \mathcal{D}_{uv}, \quad \sqrt{1 - \alpha} = \overline{\mathcal{D}}_{uv}. \quad (45)$$

Using Equations (45) and (44) in Equation (43), we get a generic form for optimal  $|\xi_i\rangle, |\zeta_j\rangle$  as in Equation (42).  $\square$

TABLE I: Values of  $P_{\lambda x}, P_{\lambda y}, G_{xy}(\lambda)$  for the general form of  $|\xi_i\rangle, |\zeta_j\rangle$  as in Equations (43).

$\lambda$	$P_{\lambda x}$	$P_{\lambda y}$	$G_{xy}(\lambda) = \frac{ P_{\lambda x} - P_{\lambda y} }{P_{\lambda x} + P_{\lambda y}}$
0	$(1 - D_{xy}) \langle \xi_x   E_0 \rangle^2 = (1 - D_{xy}) \alpha$	$(1 - D_{xy}) \langle \xi_y   E_0 \rangle^2 = (1 - D_{xy}) \beta$	$\frac{ \alpha - \beta }{\alpha + \beta}$
1	$(1 - D_{xy}) \langle \xi_x   E_1 \rangle^2 = (1 - D_{xy}) (1 - \alpha)$	$(1 - D_{xy}) \langle \xi_y   E_1 \rangle^2 = (1 - D_{xy}) (1 - \beta)$	$\frac{ \alpha - \beta }{1 - \alpha + 1 - \beta}$
2	$(D_{xy}) \langle \zeta_x   E_2 \rangle^2 = (D_{xy}) \mu$	$(D_{xy}) \langle \zeta_y   E_2 \rangle^2 = (D_{xy}) \nu$	$\frac{ \mu - \nu }{\mu + \nu}$
3	$(D_{xy}) \langle \zeta_x   E_3 \rangle^2 = (D_{xy}) (1 - \mu)$	$(D_{xy}) \langle \zeta_y   E_3 \rangle^2 = (D_{xy}) (1 - \nu)$	$\frac{ \mu - \nu }{1 - \mu + 1 - \nu}$

**Remark 2.** For equal error rates,  $\mathcal{D}_{uv}, \bar{\mathcal{D}}_{uv}$  will be replaced by  $\mathcal{D}, \bar{\mathcal{D}}$  respectively in Equation (42).

**Remark 3.** We can rewrite Equation (42) as below.

$$\begin{aligned}
|\xi_x\rangle &= \sqrt{1 - D_{uv}} |\tilde{E}_0\rangle + \sqrt{D_{uv}} |\tilde{E}_1\rangle, \\
|\xi_y\rangle &= \sqrt{1 - D_{uv}} |\tilde{E}_0\rangle - \sqrt{D_{uv}} |\tilde{E}_1\rangle, \\
|\zeta_x\rangle &= \sqrt{1 - D_{uv}} |\tilde{E}_2\rangle + \sqrt{D_{uv}} |\tilde{E}_3\rangle, \\
|\zeta_y\rangle &= \sqrt{1 - D_{uv}} |\tilde{E}_2\rangle - \sqrt{D_{uv}} |\tilde{E}_3\rangle,
\end{aligned} \quad (46)$$

where

$$\begin{aligned}
|\tilde{E}_0\rangle &= \frac{1}{\sqrt{2}} (|E_0\rangle + |E_1\rangle), & |\tilde{E}_1\rangle &= \frac{1}{\sqrt{2}} (|E_0\rangle - |E_1\rangle), \\
|\tilde{E}_2\rangle &= \frac{1}{\sqrt{2}} (|E_2\rangle + |E_3\rangle), & |\tilde{E}_3\rangle &= \frac{1}{\sqrt{2}} (|E_2\rangle - |E_3\rangle)
\end{aligned} \quad (47)$$

is another orthonormal basis (called, Bell basis), written in terms of an optimal eigenbasis  $\{E_\lambda\}$ . Clearly, these form to describe  $|\xi_i\rangle, |\zeta_j\rangle$  is analogous to Equation (50), (51) in [2].

So far, we deduced a unique representation for an optimal interaction while expressed in the eigenbasis. We postpone our task to find the corresponding optimal POVM and verification for optimality till Subsection IVD and IVE respectively. Before we go there, we need to find an expression for  $\Gamma_{xy}$  for the form of the optimal interaction that we have derived.

Since the expression (42) of  $|\xi_i\rangle, |\zeta_j\rangle$  corresponds to optimal  $G$ , therefore, by Lemma 1,  $\Gamma_{xy}$  should become a diagonal matrix for this form of  $|\xi_i\rangle, |\zeta_j\rangle$  in its eigenbasis. We confirm this, while we figure out the eigenvalues, as shown below.

**Theorem 3.** For an optimal POVM  $\{E_\lambda\}$ ,

$$\begin{aligned}
\Gamma_{xy} &= \left( \mathcal{D}_{uv}^2 - \bar{\mathcal{D}}_{uv}^2 \right) [(1 - D_{xy}) (\mathbb{E}_{00} - \mathbb{E}_{11}) \\
&\quad + D_{xy} (\mathbb{E}_{22} - \mathbb{E}_{33})],
\end{aligned} \quad (48)$$

where

$$\mathbb{E}_{ij} := |E_i\rangle\langle E_j|.$$

*Proof.* By Equation (41),

$$\Gamma_{xy} := \rho_x - \rho_y = (1 - D_{xy}) (\hat{\xi}_x - \hat{\xi}_y) + D_{xy} (\hat{\zeta}_x - \hat{\zeta}_y)$$

Using expressions of  $|\xi_i\rangle, |\zeta_j\rangle$  in Equation (42), we get,

$$\begin{aligned}
\hat{\xi}_x &= \mathcal{D}_{uv}^2 \mathbb{E}_{00} + \bar{\mathcal{D}}_{uv}^2 \mathbb{E}_{11} + 2\mathcal{D}_{uv} \cdot \bar{\mathcal{D}}_{uv} (\mathbb{E}_{01} + \mathbb{E}_{10}) \\
\hat{\xi}_y &= \bar{\mathcal{D}}_{uv}^2 \mathbb{E}_{00} + \mathcal{D}_{uv}^2 \mathbb{E}_{11} + 2\mathcal{D}_{uv} \cdot \bar{\mathcal{D}}_{uv} (\mathbb{E}_{01} + \mathbb{E}_{10}) \\
\hat{\zeta}_x &= \mathcal{D}_{uv}^2 \mathbb{E}_{22} + \bar{\mathcal{D}}_{uv}^2 \mathbb{E}_{33} + 2\mathcal{D}_{uv} \cdot \bar{\mathcal{D}}_{uv} (\mathbb{E}_{23} + \mathbb{E}_{32}) \\
\hat{\zeta}_y &= \bar{\mathcal{D}}_{uv}^2 \mathbb{E}_{22} + \mathcal{D}_{uv}^2 \mathbb{E}_{33} + 2\mathcal{D}_{uv} \cdot \bar{\mathcal{D}}_{uv} (\mathbb{E}_{23} + \mathbb{E}_{32})
\end{aligned}$$

which leads to the desired form of  $\Gamma_{xy}$ .  $\square$

**Remark 4.** Clearly, the eigenbasis diagonalizes  $\Gamma_{xy}$  with the eigenvalues

$$\begin{aligned}
\gamma_0 &= \left( \mathcal{D}_{uv}^2 - \bar{\mathcal{D}}_{uv}^2 \right) (1 - D_{xy}), & \gamma_1 &= -\gamma_0, \\
\gamma_2 &= \left( \mathcal{D}_{uv}^2 - \bar{\mathcal{D}}_{uv}^2 \right) D_{xy}, & \gamma_3 &= -\gamma_2.
\end{aligned} \quad (49)$$

Before discussing the optimal POVM for our description of optimal interaction, we consider a special case - we derive the POVM for the optimal interaction (31) for equal error rates in the next section.

### C. Optimal POVM for the Specific Interaction for Equal Error Rates ( $D_{xy} = D_{uv} = D$ ) by Fuchs et al. [2]

For equal error rates, i.e.,  $D_{xy} = D_{uv} = D$ , [2] describes a choice of  $|\xi_i\rangle, |\zeta_j\rangle$ , that optimizes  $I$  (and therefore  $G$ ). For this optimal  $|\xi_i\rangle, |\zeta_j\rangle$  as described in Equation (31), we now derive the optimal POVM that was not shown in [2].

**Theorem 4.** Consider a canonical basis for Eve as given in Equations (23). For the optimal interactions (31), the optimal POVM  $\{E_\lambda\}$  is given as

$$E_\lambda = |E_\lambda\rangle\langle E_\lambda|,$$

where

$$\begin{aligned}
|E_0\rangle &= \mathcal{D}|\mathcal{E}_0\rangle - \bar{\mathcal{D}}|\mathcal{E}_1\rangle, & |E_1\rangle &= \bar{\mathcal{D}}|\mathcal{E}_0\rangle + \mathcal{D}|\mathcal{E}_1\rangle, \\
|E_2\rangle &= \mathcal{D}|\mathcal{E}_2\rangle - \bar{\mathcal{D}}|\mathcal{E}_3\rangle, & |E_3\rangle &= \bar{\mathcal{D}}|\mathcal{E}_2\rangle + \mathcal{D}|\mathcal{E}_3\rangle.
\end{aligned} \quad (50)$$

*Proof.* Comparing a special form of  $|\xi_x\rangle, |\xi_y\rangle$  given by Equation (31) and the general form of  $|\xi_x\rangle, |\xi_y\rangle$  described in Equation (42) but for equal error rates, we get

$$\begin{aligned}\mathcal{D} |E_0\rangle + \overline{\mathcal{D}} |E_1\rangle &= |\mathcal{E}_0\rangle, \\ \overline{\mathcal{D}} |E_0\rangle + \mathcal{D} |E_1\rangle &= 2\mathcal{D} \cdot \overline{\mathcal{D}} |\mathcal{E}_0\rangle + (\mathcal{D}^2 - \overline{\mathcal{D}}^2) |\mathcal{E}_1\rangle.\end{aligned}\tag{51}$$

Solving for  $|E_0\rangle$  and  $|E_1\rangle$ , we get

$$\begin{aligned}|E_0\rangle &= \mathcal{D} |\mathcal{E}_0\rangle - \overline{\mathcal{D}} |\mathcal{E}_1\rangle, \\ |E_1\rangle &= \overline{\mathcal{D}} |\mathcal{E}_0\rangle + \mathcal{D} |\mathcal{E}_1\rangle.\end{aligned}$$

Similarly, comparing the expressions for  $|\zeta_x\rangle, |\zeta_y\rangle$  in Equations (31),(42), we can establish that

$$\begin{aligned}|E_2\rangle &= \mathcal{D} |\mathcal{E}_2\rangle - \overline{\mathcal{D}} |\mathcal{E}_3\rangle, \\ |E_3\rangle &= \overline{\mathcal{D}} |\mathcal{E}_2\rangle + \mathcal{D} |\mathcal{E}_3\rangle.\end{aligned}$$

□

We can easily check that  $\Gamma_{xy}$  is diagonalized by the eigenbasis  $\{E_\lambda\}_{\lambda \in \{0,1,2,3\}}$  given by Equation (50).

**Theorem 5.** For  $\Gamma_{xy}$  described in Equation (48), we prove that  $\Gamma_{xy}|E_0\rangle = \gamma_0|E_0\rangle$ ,

where  $\gamma_0 := 2\sqrt{D(1-D)}(1-D) = (\mathcal{D}^2 - \overline{\mathcal{D}}^2)(1-D)$  for equal error rates.

*Proof.* For equal error rates, the expression of  $\Gamma_{xy}$  in Equation (48) becomes

$$\begin{aligned}\Gamma_{xy} &= (\mathcal{D}^2 - \overline{\mathcal{D}}^2) [(1-D)(\mathbb{E}_{00} - \mathbb{E}_{11}) \\ &\quad + D(\mathbb{E}_{22} - \mathbb{E}_{33})].\end{aligned}$$

It follows immediately that

$$\Gamma_{xy}|E_0\rangle = (\mathcal{D}^2 - \overline{\mathcal{D}}^2)(1-D)|E_0\rangle = \gamma_0|E_0\rangle,$$

so far  $\{|E_i\rangle\}$  are orthonormal, which indeed is true for Equations (50). Because, in that case,

$$\mathbb{E}_{ii}|E_j\rangle = \begin{cases} |E_i\rangle, & \text{if } i = j, \\ \mathbf{0}, & \text{if } i \neq j. \end{cases}$$

This completes the proof for equal error rates, for  $\lambda = 0$ . Similarly, for unequal error rates,  $\mathcal{D}$  will be replaced by  $\mathcal{D}_{uv}$  in the above formula. □

**Remark 5.** One can calculate and check that the eigenvalues of  $\Gamma_{xy}$  described in eigenbasis (50) match those as in Equation (49) calculated for the generic form of optimal  $|\xi_i\rangle, |\zeta_j\rangle$  described by Equation (42).

This section has provided us with an insight of a possible generic form of the optimal POVM. We continue with the task to find a generic form of the optimal POVM in the next section.

#### D. Optimal POVM to maximize Information Gain (G): a Generic Form

In Theorem 5, we showed that a specific rotation of the canonical basis  $\{|\mathcal{E}_\lambda\rangle\}$  yields an eigenbasis  $\{|E_\lambda\rangle\}$  of  $\Gamma_{xy}$ . Now, we will show that not only the above specific rotation, but any rotation represented by an orthogonal linear transformation of the canonical basis  $\{|\mathcal{E}_\lambda\rangle\}$  yields an eigenbasis of  $\Gamma_{xy}$ . In the next subsection, we proceed one step further to show that such a transformation preserves the optimality of information gain as well.

**Theorem 6.** Consider a canonical basis  $\{|\mathcal{E}_\lambda\rangle\}$  for Eve as given in Equations (23). Let

$$(|E_i\rangle) = \mathbf{R}(|\mathcal{E}_j\rangle) \tag{53}$$

where  $(|E_i\rangle)$  corresponds to the column vector  $(|E_0\rangle, |E_1\rangle, |E_2\rangle, |E_3\rangle)^T$  (similar meaning for  $(|\mathcal{E}_j\rangle)$ ) and  $\mathbf{R}$  is an orthogonal matrix. Then  $\{E_\lambda\}_{\lambda \in \{0,1,2,3\}}$  forms an orthonormal eigenbasis for  $\Gamma_{xy}$ .

*Proof.* Since  $\mathbf{R}$  is an orthogonal matrix,  $\{E_\lambda\}_{\lambda \in \{0,1,2,3\}}$  are orthonormal.

For equal error rates, the expression of  $\Gamma_{xy}$  in Equation (48) becomes

$$\begin{aligned}\Gamma_{xy} &= (\mathcal{D}^2 - \overline{\mathcal{D}}^2) [(1-D)(\mathbb{E}_{00} - \mathbb{E}_{11}) \\ &\quad + D(\mathbb{E}_{22} - \mathbb{E}_{33})].\end{aligned}$$

It follows immediately that

$$\Gamma_{xy}|E_0\rangle = (\mathcal{D}^2 - \overline{\mathcal{D}}^2)(1-D)|E_0\rangle,$$

so far  $\{|E_i\rangle\}$  are orthonormal, which indeed is true for Equations (53). Because, in that case,

$$\mathbb{E}_{ii}|E_j\rangle = \begin{cases} |E_i\rangle, & \text{if } i = j, \\ \mathbf{0}, & \text{if } i \neq j. \end{cases}$$

This completes the proof for equal error rates.

Similarly, for unequal error rates,  $\mathcal{D}$  will be replaced by  $\mathcal{D}_{uv}$  in the above formula. □

Here we create a subclass of such orthogonal rotation.

**Example 1.** Consider a canonical basis  $\{|\mathcal{E}_\lambda\rangle\}$  for Eve as given in Equations (23). Let

$$\begin{aligned}|E_0\rangle &= \sqrt{a}|\mathcal{E}_0\rangle - \sqrt{1-a}|\mathcal{E}_1\rangle, \\ |E_1\rangle &= \sqrt{1-a}|\mathcal{E}_0\rangle + \sqrt{a}|\mathcal{E}_1\rangle, \\ |E_2\rangle &= \sqrt{a}|\mathcal{E}_2\rangle - \sqrt{1-a}|\mathcal{E}_3\rangle, \\ |E_3\rangle &= \sqrt{1-a}|\mathcal{E}_2\rangle + \sqrt{a}|\mathcal{E}_3\rangle.\end{aligned}\tag{54}$$

Since the coefficient matrix is orthogonal,  $\{E_\lambda\}_{\lambda \in \{0,1,2,3\}}$  forms an orthonormal eigenbasis for  $\Gamma_{xy}$ .



### E. Achieving the Optimal Information Gain (G)

Now we show that the interaction given by Equation (42) and the POVM corresponding to the eigenbasis given by Equation (53) lead to optimal  $G$ . To do so, we need to prove that they satisfy the necessary and sufficient conditions given by Proposition 3. The initial task is to find an expression for  $|\xi_u\rangle, |\xi_v\rangle, |\zeta_u\rangle, |\zeta_v\rangle$  using Equation (42). For this, we use Equations (46) and (47) in Equation (21) and (22), to derive the following intermediate result.

**Lemma 3.** *For achieving the optimal information gain, we must have*

$$\begin{aligned} |\xi_u\rangle &= \sqrt{1-D_{xy}} |\tilde{E}_0\rangle + \sqrt{D_{xy}} |\tilde{E}_2\rangle \\ |\xi_v\rangle &= \sqrt{1-D_{xy}} |\tilde{E}_0\rangle - \sqrt{D_{xy}} |\tilde{E}_2\rangle \\ |\zeta_u\rangle &= \sqrt{1-D_{xy}} |\tilde{E}_1\rangle - \sqrt{D_{xy}} |\tilde{E}_3\rangle \\ |\zeta_v\rangle &= \sqrt{1-D_{xy}} |\tilde{E}_1\rangle + \sqrt{D_{xy}} |\tilde{E}_3\rangle \end{aligned} \quad (55)$$

where the basis  $\{|\tilde{E}_\lambda\rangle\}$  is as described in Equation (47).

**Remark 6.** *To get expressions of  $|\xi_i\rangle, |\zeta_j\rangle$  in  $u$ - $v$  basis symmetric to those in  $x$ - $y$  basis, e.g., like [2, Equation (52)], one must consider the canonical basis in the order  $|\mathcal{E}_0\rangle = |u\rangle|u\rangle, |\mathcal{E}_1\rangle = |v\rangle|v\rangle, |\mathcal{E}_2\rangle = |u\rangle|v\rangle, |\mathcal{E}_3\rangle = |v\rangle|u\rangle$ , compatible with [2].*

**Theorem 7.** *The interaction given by Equation (42) and the POVM corresponding to the eigenbasis given by Equation (53) satisfy the necessary and sufficient conditions given by Proposition 3 and therefore attain optimal information gain.*

*Proof.* We have

$$\begin{aligned} |U_{\lambda u}\rangle &= B_u \otimes \sqrt{E_\lambda} |U\rangle = B_u \otimes E_\lambda |U\rangle \\ &= \sqrt{1-D_{uv}} (B_u|u\rangle) \otimes (E_\lambda|\xi_u\rangle) \\ &\quad + \sqrt{D_{uv}} (B_u|v\rangle) \otimes (E_\lambda|\zeta_v\rangle), \text{ by Eq. (20).} \end{aligned}$$

Since  $B_u|u\rangle = |u\rangle, B_u|v\rangle = \mathbf{0}$ , and  $E_\lambda|\xi_u\rangle = \langle E_\lambda|\xi_u\rangle|E_\lambda\rangle$ , we get,

$$|U_{\lambda u}\rangle = \sqrt{1-D_{uv}} \langle E_\lambda|\xi_u\rangle|u\rangle|E_\lambda\rangle.$$

Similarly,

$$|V_{\lambda u}\rangle = \sqrt{D_{uv}} \langle E_\lambda|\zeta_v\rangle|u\rangle|E_\lambda\rangle.$$

Here, we want equality in magnitude between  $\langle E_\lambda|\xi_u\rangle$  and  $\langle E_\lambda|\zeta_v\rangle$ . Now, by Eq. (55),  $\langle E_\lambda|\xi_u\rangle$  takes values  $\frac{1}{\sqrt{2}}\sqrt{1-D_{xy}}, \frac{1}{\sqrt{2}}\sqrt{1-D_{xy}}, \frac{1}{\sqrt{2}}\sqrt{D_{xy}}, \frac{1}{\sqrt{2}}\sqrt{D_{xy}}$ ; whereas,  $\langle E_\lambda|\zeta_v\rangle$  takes values  $\frac{1}{\sqrt{2}}\sqrt{1-D_{xy}}, -\frac{1}{\sqrt{2}}\sqrt{1-D_{xy}}, \frac{1}{\sqrt{2}}\sqrt{D_{xy}}, -\frac{1}{\sqrt{2}}\sqrt{D_{xy}}$ , respectively for  $\lambda = 0, 1, 2, 3$ . Therefore,

$$|V_{\lambda u}\rangle = \varepsilon_\lambda \sqrt{\frac{D_{uv}}{1-D_{uv}}} |U_{\lambda u}\rangle,$$

where,

$$\varepsilon_0 = +1, \quad \varepsilon_1 = -1, \quad \varepsilon_2 = +1, \quad \varepsilon_3 = -1. \quad (56)$$

Similarly, one may calculate to verify that

$$|U_{\lambda v}\rangle = \varepsilon_\lambda \sqrt{\frac{D_{uv}}{1-D_{uv}}} |V_{\lambda v}\rangle,$$

for the same combination of  $\varepsilon_\lambda$  as in Eq. 56. This completes the proof of the theorem.  $\square$

Further, we take the opportunity to establish a direct relation between the sign parameter  $\varepsilon_\lambda$  and the signs of eigenvalues  $\gamma_\lambda$ .

**Lemma 4.** *For optimal  $G$ ,*

$$\varepsilon_\lambda = \text{sgn } \gamma_\lambda. \quad (57)$$

*Proof.* For optimal  $G$ ,  $\Gamma$  is a diagonal matrix with diagonal entries  $\gamma_\lambda$ . Thus,

$$\gamma_\lambda = \text{tr}(\Gamma E_\lambda) = \text{tr}(\rho_x E_\lambda) - \text{tr}(\rho_y E_\lambda) = P_{\lambda x} - P_{\lambda y}.$$

By Equation (16),

$$\varepsilon_\lambda = \text{sgn}(Q_{x\lambda} - Q_{y\lambda}) = \text{sgn}(P_{\lambda x} - P_{\lambda y}) = \text{sgn } \gamma_\lambda,$$

which establishes the relation.  $\square$

**Remark 7.** *By Lemma 4, another indication for optimality is that Equation (56) should match with the signs of the eigenvalues  $\gamma_\lambda$  of  $\Gamma_{xy}$  as in Equation (49), which indeed happens here. Therefore, the optimality is achieved for the interaction given by Equation (42) and the POVM corresponding to the eigenbasis given by Equation (53).*

To summarize the results achieved so far, although there are infinitely many optimal interactions while expressed in canonical basis, they all have a unique form while written in Eigenbasis of  $\Gamma_{xy}$ . Figure 2 illustrates this fact.

## V. A DISCUSSION ON CONNECTION BETWEEN OUR RESULTS AND THOSE OF FUCHS ET AL. [2]

Here we show that the instances of an optimal interaction presented in [2] is a special case of the generalized unique form of the optimal interaction that we have derived. Moreover, we generate a new instance (different from the two instances of [2]) of the optimal interaction to add more clarity to our achievement.

As discussed in Section IV, an optimal interaction has a general form given by Equation (42). Further, the corresponding optimal POVM described in a canonical basis (23) is captured by the eigenbasis expressed in Equation (53). For a special type of the orthonormal matrix

$$\begin{aligned} |\xi_x\rangle &= \sqrt{1-D_{uv}} |\mathcal{E}_0\rangle - \sqrt{D_{uv}} |\mathcal{E}_1\rangle, \\ |\xi_y\rangle &= \sqrt{1-D_{uv}} |\mathcal{E}_0\rangle + \sqrt{D_{uv}} |\mathcal{E}_1\rangle, \\ |\zeta_x\rangle &= \sqrt{1-D_{uv}} |\mathcal{E}_2\rangle - \sqrt{D_{uv}} |\mathcal{E}_3\rangle, \\ |\zeta_y\rangle &= \sqrt{1-D_{uv}} |\mathcal{E}_2\rangle + \sqrt{D_{uv}} |\mathcal{E}_3\rangle, \end{aligned} \quad (59)$$

and the corresponding optimal POVM becomes

$$\begin{aligned} |E_0\rangle &= \frac{1}{\sqrt{2}} (|\mathcal{E}_0\rangle - |\mathcal{E}_1\rangle), & |E_1\rangle &= \frac{1}{\sqrt{2}} (|\mathcal{E}_0\rangle + |\mathcal{E}_1\rangle), \\ |E_2\rangle &= \frac{1}{\sqrt{2}} (|\mathcal{E}_2\rangle - |\mathcal{E}_3\rangle), & |E_3\rangle &= \frac{1}{\sqrt{2}} (|\mathcal{E}_2\rangle + |\mathcal{E}_3\rangle). \end{aligned}$$

One may easily check that the interaction presented here is indeed optimal. Clearly, the general form of the optimal interaction provided in this paper yields different choices of those in [2]. Moreover, its implementation is independent of equal or unequal error rates.

At this stage, we summarize the results achieved by us. Fuchs et al. [2] came up with two different configurations for optimal interactions expressed in canonical basis. For the first configuration (Equations 25), they described the corresponding POVM (Equations 26) w.r.t. the canonical basis, while for their second configuration (Equations 31), we have deduced the corresponding POVM (Equations 50) in terms of the canonical basis. We have presented one more instance of an optimal interaction (Equations 59) and the corresponding POVM (Equations 60) w.r.t. the canonical basis. For each of these three instances of the optimal interaction, one may use the relation between the eigenbasis and the canonical basis (looking at the POVM) to express the interaction w.r.t. the eigenbasis and notice that the final form becomes the same (Equations 42) for all these cases. It turns out that every possible instances of an optimal interaction written w.r.t. the canonical basis can be transformed to a unique description (Equations 42) in terms of the eigenbasis via the corresponding POVM. This is the novelty of our work. We could establish that there exists infinitely many possible instances of an optimal interaction represented in a canonical basis, but they all have a unique representation while expressed in the eigenbasis. Table II describes the general form of the optimal interaction and also shows its four specific instantiations, of which the first two coincide with those of Fuchs et al. [2] and the later two with our examples discussed earlier. The corresponding POVMs are also captured there. Figure 3 shows that feeding a POVM to the optimal form of the interaction produces a specific form of an optimal interaction.

Since an optimal interaction has a unique form and the form in Fuchs et al. [2] is a special case, any instance of such an optimal interaction will achieve the same optimal information gain  $G^*$  benchmarked in [2], neither more nor less. Thus, we do not claim any improvement on the optimal information gain here.

## VI. CONCLUSION

For the BB84 quantum key distribution protocol, we have established a unique form describing the optimal interaction followed by the corresponding optimal measurement for the optimal information gain an eavesdropper can obtain for a given average disturbance when her

Optimal Interaction	Ours: General [Eq. 42]	Fuchs 1 [Eq. 27]	Fuchs 2 [Eq. 31]	Ours: One Parameter [Eq. 58]	Ours: Special Case [Eq. 59]
$ \xi_x\rangle$	$\mathcal{D}_{uv}  E_0\rangle + \overline{\mathcal{D}}_{uv}  E_1\rangle$	$\mathcal{D}_{uv}  \mathcal{E}_0\rangle + \overline{\mathcal{D}}_{uv}  \mathcal{E}_3\rangle$	$ \mathcal{E}_0\rangle$	$(\mathcal{D}_{uv}\sqrt{a} + \overline{\mathcal{D}}_{uv}\sqrt{1-a})  \mathcal{E}_0\rangle + (\overline{\mathcal{D}}_{uv}\sqrt{a} - \mathcal{D}_{uv}\sqrt{1-a})  \mathcal{E}_1\rangle$	$\sqrt{1-D_{uv}}  \mathcal{E}_0\rangle - \sqrt{D_{uv}}  \mathcal{E}_1\rangle$
$ \xi_y\rangle$	$\overline{\mathcal{D}}_{uv}  E_0\rangle + \mathcal{D}_{uv}  E_1\rangle$	$\overline{\mathcal{D}}_{uv}  \mathcal{E}_0\rangle + \mathcal{D}_{uv}  \mathcal{E}_3\rangle$	$2\mathcal{D}\overline{\mathcal{D}}  \mathcal{E}_0\rangle + (\mathcal{D}^2 - \overline{\mathcal{D}}^2)  \mathcal{E}_1\rangle$	$(\overline{\mathcal{D}}_{uv}\sqrt{a} + \mathcal{D}_{uv}\sqrt{1-a})  \mathcal{E}_0\rangle + (\mathcal{D}_{uv}\sqrt{a} - \overline{\mathcal{D}}_{uv}\sqrt{1-a})  \mathcal{E}_1\rangle$	$\sqrt{1-D_{uv}}  \mathcal{E}_0\rangle + \sqrt{D_{uv}}  \mathcal{E}_1\rangle$
$ \zeta_x\rangle$	$\mathcal{D}_{uv}  E_2\rangle + \overline{\mathcal{D}}_{uv}  E_3\rangle$	$\overline{\mathcal{D}}_{uv}  \mathcal{E}_2\rangle + \mathcal{D}_{uv}  \mathcal{E}_1\rangle$	$ \mathcal{E}_2\rangle$	$(\mathcal{D}_{uv}\sqrt{a} + \overline{\mathcal{D}}_{uv}\sqrt{1-a})  \mathcal{E}_2\rangle + (\overline{\mathcal{D}}_{uv}\sqrt{a} - \mathcal{D}_{uv}\sqrt{1-a})  \mathcal{E}_3\rangle$	$\sqrt{1-D_{uv}}  \mathcal{E}_2\rangle - \sqrt{D_{uv}}  \mathcal{E}_3\rangle$
$ \zeta_y\rangle$	$\overline{\mathcal{D}}_{uv}  E_2\rangle + \mathcal{D}_{uv}  E_3\rangle$	$\mathcal{D}_{uv}  \mathcal{E}_2\rangle + \overline{\mathcal{D}}_{uv}  \mathcal{E}_1\rangle$	$2\mathcal{D}\overline{\mathcal{D}}  \mathcal{E}_2\rangle + (\mathcal{D}^2 - \overline{\mathcal{D}}^2)  \mathcal{E}_3\rangle$	$(\overline{\mathcal{D}}_{uv}\sqrt{a} + \mathcal{D}_{uv}\sqrt{1-a})  \mathcal{E}_2\rangle + (\mathcal{D}_{uv}\sqrt{a} - \overline{\mathcal{D}}_{uv}\sqrt{1-a})  \mathcal{E}_3\rangle$	$\sqrt{1-D_{uv}}  \mathcal{E}_2\rangle + \sqrt{D_{uv}}  \mathcal{E}_3\rangle$
POVM	Ours: General	Fuchs 1 [Eq. 26]	Fuchs 2 [Eq. 50]	Ours: One Parameter [Eq. 54]	Ours: Special Case [Eq. 60]
$ E_0\rangle$	$ E_0\rangle$	$ \mathcal{E}_0\rangle$	$\mathcal{D} \mathcal{E}_0\rangle - \overline{\mathcal{D}} \mathcal{E}_1\rangle$	$\sqrt{a} \mathcal{E}_0\rangle - \sqrt{1-a} \mathcal{E}_1\rangle$	$\frac{1}{\sqrt{2}} ( \mathcal{E}_0\rangle -  \mathcal{E}_1\rangle)$
$ E_1\rangle$	$ E_1\rangle$	$ \mathcal{E}_1\rangle$	$\overline{\mathcal{D}} \mathcal{E}_0\rangle + \mathcal{D} \mathcal{E}_1\rangle$	$\sqrt{1-a} \mathcal{E}_0\rangle + \sqrt{a} \mathcal{E}_1\rangle$	$\frac{1}{\sqrt{2}} ( \mathcal{E}_0\rangle +  \mathcal{E}_1\rangle)$
$ E_2\rangle$	$ E_2\rangle$	$ \mathcal{E}_2\rangle$	$\mathcal{D} \mathcal{E}_2\rangle - \overline{\mathcal{D}} \mathcal{E}_3\rangle$	$\sqrt{a} \mathcal{E}_2\rangle - \sqrt{1-a} \mathcal{E}_3\rangle$	$\frac{1}{\sqrt{2}} ( \mathcal{E}_2\rangle -  \mathcal{E}_3\rangle)$
$ E_3\rangle$	$ E_3\rangle$	$ \mathcal{E}_3\rangle$	$\overline{\mathcal{D}} \mathcal{E}_2\rangle + \mathcal{D} \mathcal{E}_3\rangle$	$\sqrt{1-a} \mathcal{E}_2\rangle + \sqrt{a} \mathcal{E}_3\rangle$	$\frac{1}{\sqrt{2}} ( \mathcal{E}_2\rangle +  \mathcal{E}_3\rangle)$

TABLE II: Optimal interaction and corresponding POVM - unique general form and its specific instances.

interaction and measurements are performed signal by signal. We have shown that the choice of optimal inter-

action in [2], for equal as well as unequal error rates, is a special case of the form provided by us.

- 
- [1] C. H. Bennett and G. Brassard. *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing*, Bangalore, India, 175–179, IEEE, New York (1984).
  - [2] C. A. Fuchs, N. Gisin, R. B. Griffiths, C. S. Niu, and A. Peres. *Phys. Rev. A* **56**, 1163–1172 (1997).
  - [3] C. A. Fuchs. *Fourth Workshop on Physics and Computation* — *PHYSCOMP '96*, 229–259 (1996).
  - [4] M. A. Nielsen and I. L. Chuang. *Cambridge University Press* (2002).
  - [5] T. Cover and J. Thomas. *John Wiley & Sons, Inc.*, First Edition, 16–20 (1991).